

Chromosome 13

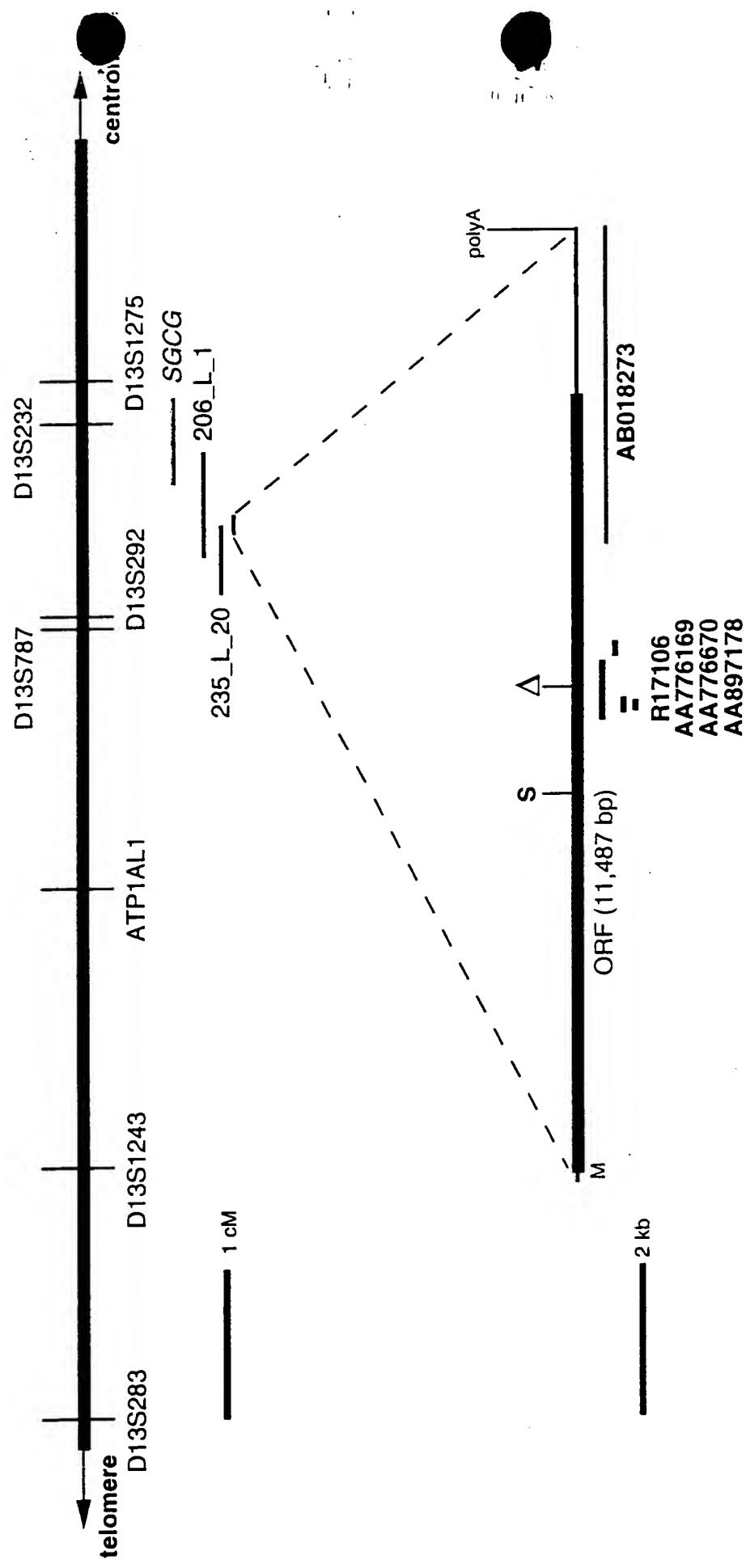
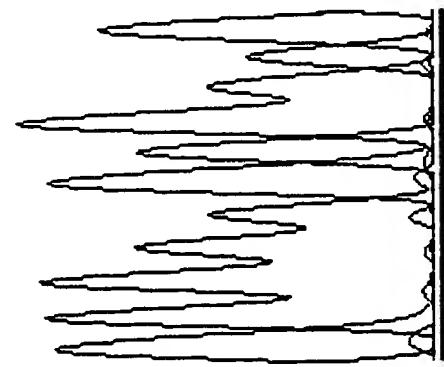


Figure 1

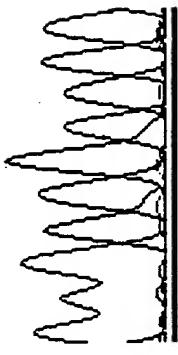
.....

A C C C T A T C A  
(SEQ ID NO: 17) ▼



6594T

A A A G C G A C A C  
(SEQ ID NO: 19) ▼

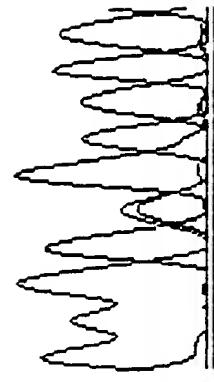


5254C

A A A G N G A C C  
(SEQ ID NO: 20) ▼



6594Δ



5254C/T

Fig. 2A

Fig. 2B

1000 900 800 700 600 500 400 300 200 100 0

M 1 2 3 4 5 6 M 7 8 9 10 11 12 13 14

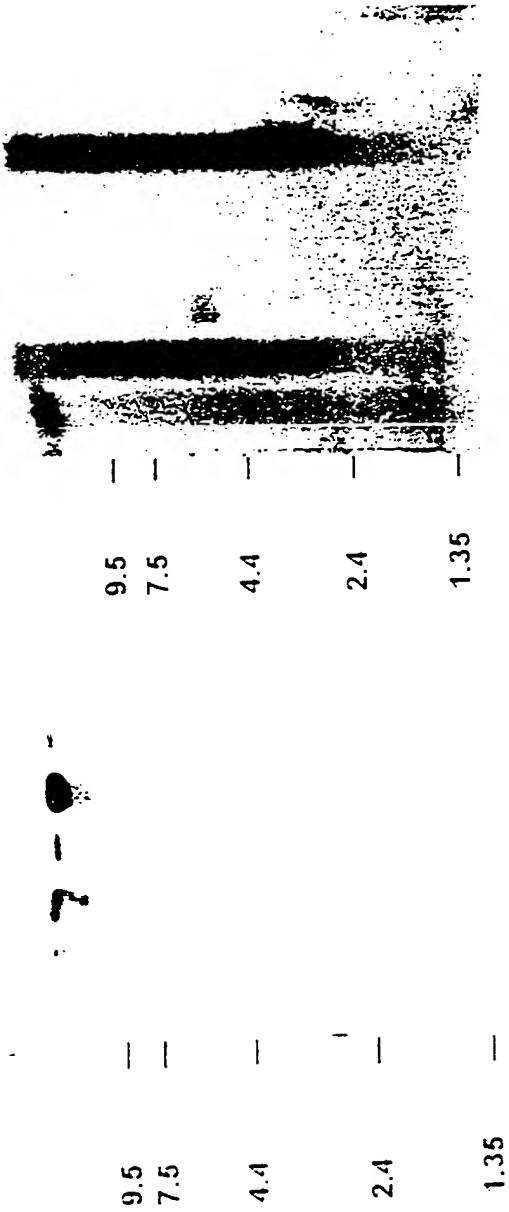


Figure 3

Fig. 4A

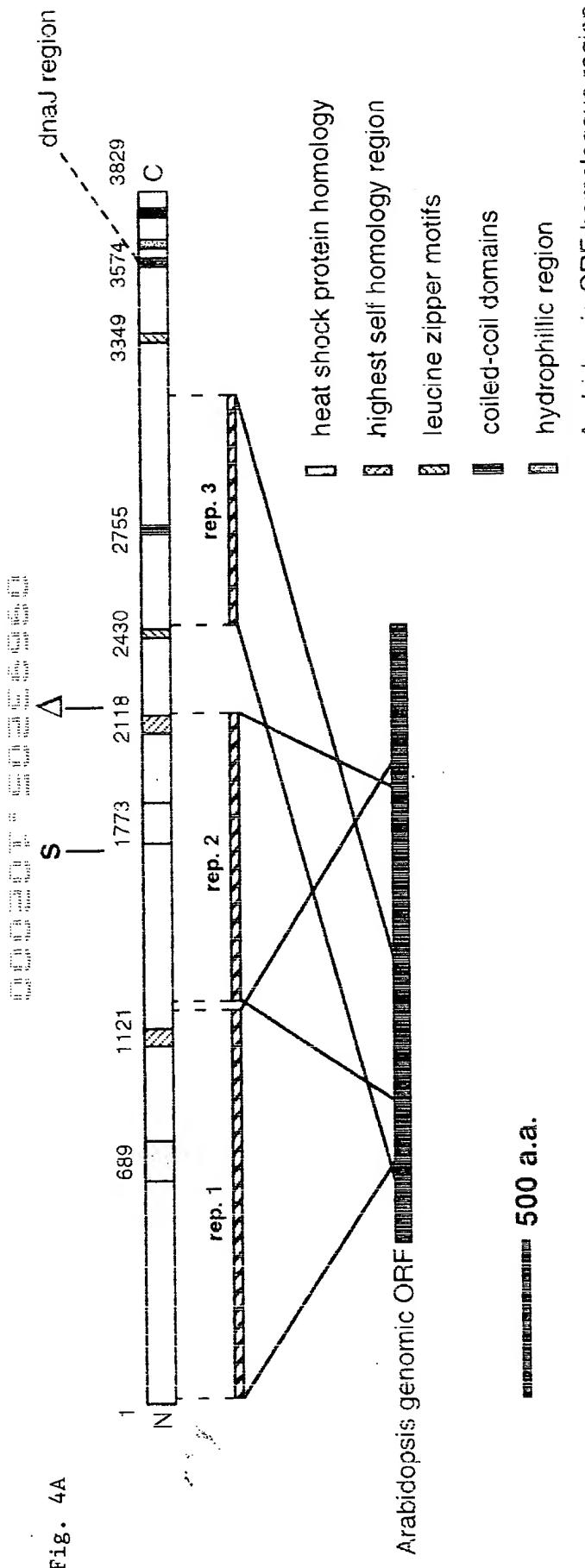


Fig. 4B

(SEQ ID NO: 2) MNTFWPGRELIVQWYPDFENRNHPSVSWL KMWNKNLVHFSEDLTLPDEMLIPRTILEE 60  
 (SEQ ID NO: 4) .....S.DKR....L.....L.N.

GQTCVELIRLRIPSLVILDDESEAQLPEFLADIVQKLGGFVLKKLDASIQHPLIKKYIHS 120  
 D.....V.....T.....I...R..T.....V.....

PLPSAVLQIMEKMPPLQKLCNQITSLLPTHKDALKFLASLTDSEKEKRIIQELAIFKRI 180  
 ....I.....I.....A.....T.....T.....T.....

NHSSDQGISSYTKLKGCKVLHHTAKLPADLRLSISVIDSSDEATIRLANMLKIEQLKTT 240  
 .....D.....T.....V.....K.....

CLKLVLKDIENAFYSHEEVTLQMLWVLENLSSLKNENPNVLEWLTPLKFIQISQEQMVA 300  
 ...F.....G....TQ.....I.....S...D...M....HM..GHV.A.

GELFDPDIEVLKDLFCNEEGTYFPSSVFTSPDILHSLRQIGLKNEASLKEKDVVQVAKKI 360  
 .D.....R...Y...EAC...TI.....S.....R..

EALQVGACPQDVLLKKAKTLLVLNKNHTLLQSSEGKMTLKKIKWVPACKERPPNPYGS 420  
 .....SS.QN....M.....Q.....A.....

LVWKGDLCNLCAAPPDMCDVGHAILEDIGSSLPLVESIHVNLEKALGIFTKPSLSAVLKHFKI 480  
 .....AA..V.V.....V.....Q..S.....TIN.....T

VVDWYSSKTFSDDEDYYQFQHILLEIYGMHDHLNEGKDSFRALKFPWWTGKKFCPLAQ 540  
 .....T.....S.....K.....N.....

VIKPIHDLDLQPYLHNVPKTMAKFHQFKCGSIEELTSDHISMVIQKLYLKSDQDLSEQ 600  
 ....T.....Y.....A.....V.....E...E

ESKQNLHMLNIIRWLYSNQIPASPNTPVPIHHSKNPSKLIMKPIHECCYCDIKVDDLND 660  
 .....M.....Y..R.....V.....

LLEDSVEPIILVHEDI PMKTAEWLKVPCESTRLINPENMGFEQSGQREPLTVRIKNILEE 720  
 .....  

YPSVSDFKELLQNADDANATECSFLIDMRRNMDIRENLLDPGMAACHGPALWSFNNSQE 780  
 .....M.....E.....

SDSDFVNITRLGESLKRGEVDKVGKFGFGLGFNSVYHITDIPIMSREFMIMFDPNINHISK 840  
 .....D.....B.....

HIKDKSNPGIKINWSKQQKRLRKFPNQFKPFIDVFGCQLPLTVEAPSYNGTLFRSFRT 900  
 ....R.....A.....

QQEAKVSEVSSTCYNTADIYSLVDEFSLCGHRLIIIFTQSVKSMYLKYLKIEETNPSLAQD 960  
 .....N.....

TVIICKKSCSSSKALNTPVLSVLKEAKLMKTCSSSNKLPSSDEPKSSCILQITVEEFHHV 1020  
 .I.....V.P.....A.....T.V.....

FRRIAQLQSPLFRGPDPPAALFEMAKSGQSKKPSDELSQLTVECTWLLCTCMĐTGEAL 1080  
 .....T.....P....D.....I.....

KFSLSESGRRLGLVPCGAVGVLSEIQDQKWTVKPHIGEVFCYPLRIKTGLPVHNGCF 1140  
 .....N.....L.H.T.E.....I.....

AVTSNRKEIWKTDTKGRWNTTFMRHVIVKAYLQVLSVLRLATSGELMDYTYAVWPDPD 1200  
 .....A.....IG....T.....

LVHDDFSVICQGFYEDIAHGKGKELTKVFDGSTWVSMKNVRFLDDSIKRRDVGSAAFK 1260  
 .....K.....R.....M.....Q.K.....

IFLKYLKKTGSKNLCAVELPSSVKLGFEAGCKQILLENTFSEKQFFSEVFFPNIQEIEA 1320  
 .....A.....

ELRDPLMIFVNLNEKVDEFSGVLRVTPCIPCSLEGHPLVLPSSLIHPEGRVAKLFDIKDGR 1380  
 .....N.....L.....I.....V.....T.....

FPYGSTQDYLNPIILIKLVLQLGAKDDILWDDMLERAVSVAEINKSDHVAACLRSSILLS 1440  
 .....M.....E.....A.....

Figure 5A

LIDEKLKIRDPRAKDFAAKYQTIRFLPFLTKPAGFSLDWKNSFKPETMFRATDLYTAEH 1500  
.....K.....P.....E.....I.....Y

QDIVCILPINENSHSFRGCGSVSLAVKEFLGLKKPTVDEQLKEVAKSVDDGITL 1560  
.....Q.....

YQENITNACYKYLHEALMQNEITKMSIIDKLKPFSFILVENAYVDSEKVSFHLNFEAAPY 1620  
.....VL.....MA.....AT.....E.....C.....V.....E.....

LYQLPNKYKNNFRELFFETVGVRQSCTVEDFALVLESIDQERGKQITEENFQLCRRIISE 1680  
.....S.....F.....K.....

GIWSLIREKKQEFCENYGKILLPDTNMLLPAKSLCYNDCPWIKVKDTTVKYCHADIPR 1740  
.....R.....L.....S.....

EVAVKLGAVPKRHKALERASNVCFITLGTEFGQKEKLTSRIKSILNAYPSEKEMLKELL 1800  
.....I.....I.....A.....

ONADDAAKATEICEVFDPROHPVDRIFDDKWAPLOGPALCVYNNOPFTEDDVRGIONLGKG 1860  
.....

TKEGNFYKIGQYGYGIGFNSVYHITDCPSFISGNDILCLFDPHARYAPCATSISPGRMFRL 1920  
.....C.....H.....G.....V.....

DADFRTOFSVDLTLGLGTHFKLDNCTMFRPLRNAEMAKVSEISSVPASDRMVQNLLDKL 1980  
.....N.....Q.....S.....

RSDGAELLMFLNHMEKISICEIDKSTGALNVLYSVKGKITDGDRLKRKQFHASVIDSVTK 2040  
.....A.....G.....

KRQLKDIKVQQITYTMDTEDSEGNLTTLICNRSGFSSMEKVSKSVISAHKNQDITLFPR 2100  
.....

GGVAACITHNYKKPHRAFCFLPLSLETGLPFHVNGHFALDSARRNLWRDDNGVGVRSDWN 2160  
.....

NSLMTALIAPAYVELLIQLKKRYFPGSDPTLSVLQNTPIHVVKDTLKKFLSFFPVNRLDL 2220  
.....

QPDLYCLVKALYNCIHEDMKRLLPVVRAPNIDGSDLHSAVIITWINMSTSNKTRPFFDNL 2280  
.....S.....

LQDELQHLKNADYNITTRKTVAENVYRLKHLLIEIGFNLVYNCDETANLYHCLIDADIPV 2340  
.....V.....

SYVTPADIRSFLMTFPDTNCHIGKLPTRLQQTNLKFHSLKLLVDYCFKDAEENEIEV 2400  
.....V.....S.F..

EGLPLLIITLDSVLQTFDAKRPKFLTTYHELI PSRKDLFMNTLYLKYSNILLCKVAKVFD 2460  
.....I.G.....SV.....

ISSFADLLSSVLPREYKTKSCTWKDNFASESWLKNAWHFISESVVKEDQEETKPTFDI 2520  
.....N.A.....TD.....P.A..V

VVDTLKD WALLPGTKFTVSANQLVVPEGDVLLPLSLMHI AVFPNAQSDKVFHALMKAGCI 2580  
I..I.....TS.....I.....

QLALNKICSKDSAFVPLLSCHTANIESPTSILKALHYMVTSTFRAEKLVENDFEALLMY 2640  
.....L.....D.A.....V.....T.....M.....

FNCNLNHLMSQDDIKILKSLPCYKSISGRYVSIGKFGTCYVLTKSIPS AVEEKWTQSSSS 2700  
.....S.....M.A.....

AFLEEKIHLKELYEVIGCVPVDDLEVYLKHLLPKIENLSYDAKLEHLIYLNRL SSAEEL 2760  
.....V.....L.....

SEIKEOLFEEKLESLLIHDANSRLKQAKHFYDRTVRVFEVMLPEKLFIPNDFKKLEQLI 2820  
.....N.....KE.....V.....

KPKNHVTFMTSWVEFLRNIGLKYI LSQQQLLQFAKEISVRANTENWSKETLQNTVDILLH 2880  
.....QAA.....A.....S.....

HIFQERMDLLSGNFLKELSLIPFLCPERAPAEFIRFHPQYQEVRNGTLPLIKFNGAQVNPK 2940  
.....Y.....

FKQCDVLQLLWTSCPILPEKATPLSIKEQEGSDLGPQEQLEQVNLNVNLDPPLDKVIN 3000  
.....A.....

NCRNICNITTLDEEMVKTRAKVLRSIYEFLSAEKREFRFQLRGVAFVMVEDGWKLLKPEE 3060  
.....

VVINLEYESDFKPVLYKLPLELGTFHQLFKHLGTEDIISTKQYVEVLSRIFKNSEGKQLD 3120  
.....A.....S.....

PNEMRTVKRVVSGLFRSLQNDSVKVRSDLENVRDLALYLPSQDGRLVKSSILVFDDAPHY 3180  
.....K.....A.....K.....

KSRIQGNIGVQMLVDLSQCYLGKDHFHTKLIMLFHQKLRPRLSSILEEQLDEETPKVC 3240  
.....

QFGALCSLQGRLQLLSSSEQFITGLIRIMKHENDNAFLANEEKAIRLCKALREGLKVS CF 3300  
.....

EKLQTTLRVKGFNPIPHSRSETFAFLKRGNAVILLYIQHSDSKDINFLLALAMTLKSAT 3360  
.....

DNLISDTSYLIAMLGCNDIYRIGEKLDLSLGVKYDSSEPSKLELPMPGTPIPAEIFYTLLM 3420  
.....S.....

DPMNVYPGEYVGYLVDAAEGGDIYGSYQPTYTYAIIVQEVEREDADNSSFLGKIYQIDIG 3480  
.....T.....

YSEYKIVSSLPLYKFSRPEESSQRDSAPSTPTSPTEFLTPGLRSIPPLFSGRESHKT-S 3540  
.....D....N....T.....K....SP.

SKHQSPKKLVNSLPEILKEVTSVVEQAWKI [REDACTED] 3600  
T..H..R.....A.....

[REDACTED] STSASRFQSDKYSFQRFYT [REDACTED] SWNOEATSHK 3660  
[REDACTED]

SERCCONKEKCPSAGQTYSQRFVPPTFKSVGNPVEARRWLRQARANFSAARNDLHKNA 3720  
.....S.....

NEWVCFKCYLSTKLALIAADYAVRGKSDKDVKP TALAOKEEYSOOLEGLTNVDVHTLEY 3780  
.....

GVDSLKTRYPDLLPFPQIPNDRFTSEVAMRVMECTACIIKLENFMQQKV 3830  
.....I.....

Figure 5C

**Table 1 ESTs identified by sample-sequencing of the ARSACS critical interval**

BAC clone	GenBank #	UniGene	Identity <sup>a</sup>	Tissue Source
235_1_20	AA987300	Hs.129092	221/230	neuroendocrine lung carcinoids
235_1_20	AA476635*		249/296	total fetus
235_1_20	AA1351876*		272/335	melanocyte, fetal heart, pregnant uterus (pool)
235_1_20	W25994	Hs.163732	447/464	retina
235_1_20	AA1377467	Hs.163732	257/263	total fetus
235_1_20	AA601007		488/491	schwannoma tumor
235_1_20	AA324964		214/234	cerebellum
235_1_20, 206_1_1	AA897178 <sup>b</sup>		238/238	fetal lung, testis, and B-cell
235_1_20, 206_1_1	R17106 <sup>b</sup>	Hs.188560	747/784	brain, adipose tissue
235_1_20, 206_1_1	AB018273 <sup>b</sup>	Hs.159492	4318/4318	multiple tissue types including brain, CNS, and whole embryo

a: Number of homologous nucleotides between the BAC subclone sequence and the GenBank sequence

b: AA897178, R17106, and AB018273 are all contained within *spastin*.

\* AA476635 and AA1351876 were homologous to non-overlapping portions of the same M13 subclone sequence.

Figure 6

**Table 2 Primers for PCR amplification of the human *spastin* ORF**

Primer set	Forward primer	Reverse primer	Product size (bp)
1 (SEQ ID NO: 21)	CCTTCCAGTACTGGTTATITGTCAG	CAAGAACCTCCTCAGGCCATC	(SEQ ID NO: 22) 603
2 (SEQ ID NO: 23)	GATGCATCTATAACATCCGGCT	GGGTGGGAATAGGTTCCTTC	(SEQ ID NO: 24) 581
3 (SEQ ID NO: 25)	AAAAATGAGAATTCCAATGTGCT	GCACTAAGGCTAGGTTTGTGAAG	(SEQ ID NO: 26) 592
4 (SEQ ID NO: 27)	GCTCCCTCACITCCCCTTGTTG	CCTGAAATTGGCTTCACTGATAA	(SEQ ID NO: 28) 602
5 (SEQ ID NO: 29)	AGCAATCAGATTCCAGCAAGC	GATGGGAATGTCACTGGTAATGC	(SEQ ID NO: 30) 611
6 (SEQ ID NO: 31)	GGGAGAAGTTGACAACAGTTGGA	CTTGGGTTCACTCACTGGGAAG	(SEQ ID NO: 32) 624
7 (SEQ ID NO: 33)	TCCAAGGCAATTGAAACACACCT	CAGGTCCCGTAAGACACTCAG	(SEQ ID NO: 34) 631
8 (SEQ ID NO: 35)	CAATGGGGCTTGCTGTITAC	CGAAGAACCTCCCGAGAACATCA	(SEQ ID NO: 36) 620
9 (SEQ ID NO: 37)	GCTGGCTGCAAACAGATACTAC	GCAAAACATGGRTICAGGCTTA	(SEQ ID NO: 38) 604
10 (SEQ ID NO: 39)	CAAACAACTCCGCCTCCCTCCAT	ATTATTGGTGGCCTAAAGCTGA	(SEQ ID NO: 40) 651
11 (SEQ ID NO: 41)	TTCCGGGAACCTTTTGAAACC	ACACAAAGTGTGGCCCTTGC	(SEQ ID NO: 42) 625
12 (SEQ ID NO: 43)	GATGCAAAGGGACAGAAATC	ATACAGCACATTAGAGCTCCAGT	(SEQ ID NO: 44) 626
13 (SEQ ID NO: 45)	GCATCAGACAGAATGGTCCAG	GCAATTCAACATATGCAGGGAG	(SEQ ID NO: 46) 624
14 (SEQ ID NO: 47)	GTGAATGGCCACTTTGCACT	TGATATCAGCAGGGGTACAT	(SEQ ID NO: 48) 648
15 (SEQ ID NO: 49)	ACACACGGAAAACAGTAGCA	GCCATGGCATTCCTLAAGCCAAAG	(SEQ ID NO: 50) 609
16 (SEQ ID NO: 51)	TGACATTTCAGCTTGTGA	AGGGGCACCTGATGGATTAT	(SEQ ID NO: 52) 631
17 (SEQ ID NO: 53)	AAATGATTTGAGGGCACTTGT	TTCCACCCAGGAATGTCAATAAA	(SEQ ID NO: 54) 609
18 (SEQ ID NO: 55)	ACAGTAGACTAAAGCAAGCAAAGC	ATCAAAGGGAGGAAACCAAGGTT	(SEQ ID NO: 56) 645
19 (SEQ ID NO: 57)	CATCCCTGCCCTATTCTTCAG	TAAGGGCAAGGTCCTGTACA	(SEQ ID NO: 58) 618
20 (SEQ ID NO: 59)	TGAGGGCAAAACAAATTAGATCC	TCTGCTGTTGGGAATAGGATT	(SEQ ID NO: 60) 612
21 (SEQ ID NO: 61)	GCAAAGCCCTAAAGAGAAGGATT	TGCTTGGAGACGCTTCCCTCAG	(SEQ ID NO: 62) 647
22 (SEQ ID NO: 63)	TGAAAGAGAAGATGCTGACAATT	GTAAGTCTGTCCGGCTGAAGG	(SEQ ID NO: 64) 654
23 (SEQ ID NO: 66)	CATCCCCGATTTCAGTCAGACA	TTCGTGCTACAAACACATCAAGA	(SEQ ID NO: 66) 638

Figure 7

LOCUS AF193557 11493 bp DNA ROD  
 DEFINITION Mus musculus sacsin gene, complete cds.  
 ACCESSION AF193557  
 VERSION AF193557.1 GI:6907043  
 KEYWORDS .  
 SOURCE house mouse.  
 ORGANISM **Mus musculus**  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 11493)  
 REFERENCE AUTHORS Engert,J.C., Berube,P., Mercier,J., Dore,C., Lepage,P., Ge,B.,  
 Bouchard,J.P., Mathieu,J., Melancon,S.B., Schalling,M.,  
 Lander,E.S., Morgan,K., Hudson,T.J. and Richter,A.  
 TITLE ARSACS, a spastic ataxia common in northeastern Quebec, is caused  
 by mutations in a new gene encoding an 11.5-kb ORF  
 JOURNAL Nat. Genet. 24 (2), 120-125 (2000)  
 MEDLINE 20120709  
 REFERENCE AUTHORS Engert,J.C., Berube,P., Dore,C., Lepage,P., Ge,B., Hudson,T.J. and  
 Richter,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-OCT-1999) Genome Centre, Montreal General Hospital,  
 1650 Cedar Ave., Montreal, QC H3G 1A4, Canada  
 FEATURES Location/Qualifiers  
 source 1..11493  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 mRNA <1..>11493  
 /product="sacsin"  
 CDS 1..11493  
 /note="molecular chaperone"  
 /codon\_start=1  
 /product="sacsin"  
 /protein\_id="**AAF31263.1**"  
 /db\_xref="GI:6907044"  
  
 /translation="MNTFWPGRELVVQWYPFSEDKRHPSLSWLKMVKNLYIHFSEDL  
 TLFDEMPLIPRTLLNEDQTCVELIRLRIPSVILDDTEAQLPEFLADIVQKLGIVL  
 KRLDTSIQHPLVKKYIHSPPLPSAILQIMEKIPLQKLCNQIASLLPTHKDALRKFLASL  
 TDTSEKEKRIIQELTIFKRINHSSDQGISSYTKLGCKVLDHTAKLPTDLRLSVSVID  
 SSDEATIRLANMLKIEKLKTTSCLKFVLKDIGNAFYTQEEVTQLMLWILENLSSLKNE  
 NSNVLDWLMPLKFIHMSQGHVVAAGDLFDPDIEVLRDLFYNEEEACFPPTIFTSPDIL  
 HSLRQIGLKNESSLKEKDVVQVARKEALQVSSCQNQDVLMKKAKTLLLVLNKNQTLL  
 QSSEGKMAKKIKWVPACKERPPNYPGSLVWKGDLCNLCAAPPDMCDAAHAVLVGSSL  
 LVESVHVNL\_EQALSIFTKPTINAVALKFKTVDWYTSKTFSDEDYYQFQHILLEIYGF  
 MHDHLSEGKDSFKALKFPWWWTGKNFCPLAQAVIKPTHDLQPYLYNVPKTMKFHQ  
 LFKACGSIEELTSDHISMVIQKVYLKSDQELSEEESKQNLHMLNIMRWLYSNQIPAS

Figure 8A

PNTPVPIYHSRNPSKLMKPIHECCYCDIKVDDLNDLLEDSVEPIILVHEDIPMKTAE  
WLKVPCLOUDLINPENMGFEQSGQREPLTVRIKNILEEYPSVSDIFKELLQNADDANA  
TECSFMIDMRRNMDIRENLLDPGMAACHGPALWSFNNSEFSDFLNITRLGESLKRG  
EVDKVGKFGLGNSVYHITDIPIMSREFMIMFDPNINHISKHIKDRSNPGIKINWSK  
QQKRLRKFPNQFKPFIDVFGCQLPLAVEAPYSYNGTLFRLSFRTQQEAKVSEVSSTCY  
NTADIYSLVDEFSLCGHRLIIFTQSVNSMYLKYLKIEETNPSLAQDTIIKKKVCPSK  
ALNAPVLSVLKEAAKLMKTCSSSNKKLPTDVKSSCILQITVEEFHHFRRRIADLQSP  
LFRGPDDDPATLFEMAKSGQSCKPSDELPQKTVDCCTWLICTCMDTGEALKFSLNESG  
RRLGLVPCGAVGVLLHETQEOKWTVKPHIGEVFCYLPRLIKTGLPIHINGCFAVTSNR  
KEIWKTDTKGRWNTTFMRHVIVKAYLQALSVLRDLAIGGELTDYTYAVWPDPDLVHD  
DFSVICKGFYEDIAHGKGKELTRVFSGDGSMWVSMKNVRFLDDSILORKDVGSAAFKIF  
LKYLKKTGSKNLCAVELPSSVKAGFEEAGCKQILLENTFSEKQFFSEVFFPNIQEIEA  
ELRDPLMNVLNEKLDEFSGILRVTPCVPCSLEGHPLVLPTRLIHPEGRVAKLFDTKD  
GRFPYGSTQDYLNPILLIKLVQLGMAKDDILWDDMLERAESVAEINKSDHAAACLRSS  
ILLSLIDEKLKIKDPRAKDFAAKYQTIPFLPFLTKPAGFSLEWKNSFKPETMFAATD  
IYTAEYQDIVCLLQPILNENSHSFRGCGSVSLAVKEFLGLKKPTVDLVINQLKQVAK  
SVDDGITLYQENITNACYKYLHEAVLQNEMAKATIEKLKPFCFILVENVYVESEKVS  
FHNLNFEAAPYLYQLPNKYKNNFRELFESEVGVRQSFTVEDFALVLESIDQERGKKQITE  
ENFQLCRRIISEGIWSLIREKRQEFCENKGILLPDTNLLLPAKSLCYNDCPWIKV  
KDSTVKYCHADIPREVAVKLGAIPKRHKALERYASNICFTALGTEFGQKEKLTSRIKS  
ILNAYPSEKEMLKELLQNADDAKATEICVFDPQRHPVDRIFDDKWAPLQGPALCVYN  
NQPFTEDDVRGIQNLGKGTKEGNPCKTGHYGIGFNSVYHITDCPSFISGNDILGIFDP  
HARYAPGATSVSPGRMFRLDADFRQFSDVLDLYLGNHFKLDNCTMFRPPLRNAEMA  
QVSEISSVPSSDRMVQNLLDKLRSRGAE MLMFLNHMEKISICEIDKATGGLNVLYSVK  
GKITDGDRLLRKQFHASVIDSVTKKRQLKDIPVQQITYTMDTEDSEGNLTWLCNRS  
GFSSMEKVSKSVISAHKNQDITLFPRGGVAACITHNYKKPHRAFCFLPLSLETGLPFH  
VNIGHFALDSARRNLWRDDNGVGVRSDWNNSLMTALIAPAYVELLIQLKKRYFPGSDPT  
LSVLQNTPIHVVKDTLKKFLSFFPVNRLDLQPDLYCLVKALYSCIHEDMKRLLPVRA

Figure 8B

PNIDGSDLHSAVIITWINMSTSINKTRPFFDNLLQDELQHLKNADYNITTRKTVAENVY  
RLKHLLLEIGFNLVYNCDETANLYHCLVDADIPVSYVTPADVRSFLMTFSSPDTNCHI  
GKLPTRLQQTNLKLHFSLKLLVDYCFKDAEESEFEVEGLPLLITLDSVLQIFDGKRPK  
FLTTYHELI PSRKDLFMNTLYLKYSVLLNCVKAVFDI SSFADLSSVLPREYTKN  
CAWKDNFASESWLKN AWHFISESVS VTDQEEP KPAFDVIVDILKDWALLPGTKFTV  
STSQLVVPEGDVLIPLSLMHIAVFPNAQSDKVFHALMKAGCIQLALNKICSKDSALVP  
LLSCHTANIDSPASILKAVHYMVQTSTFRTEKL MENDFE ALLMYFNCNLSHLMSQDDI  
KILKSLPCYKSISGRYMSIAKFGTCYVLTKSIPS AVEKWTQSSSAFLEEKVHLKEL  
YEVLGCVPDDLEVYLKHL PKIENL SYDAKLEH LIYLN RLASIEEPSEIKEQLFEK  
LESLLIIHDANNRLKQAKHFYDRTVRVFEVMLPEKLFIPKEFFKLEQVIKPKNQAAF  
MTSWVEFLRNIGLKYALSQQQLLQFAKEISVRANTENWSKETLQSTVDILLHHIFQER  
MDLLSGNFLKELSLIPFLC PERAPAEYIRFHPQYQE VNGLPLIKFNGAQVNPKFKQC  
DVLQLLWTSCPILPEKATPLSIKEQEGSDLAPQE QLEQV LMLNVNLD PPLDKVINNC  
RNICNITTLDEEMVKTRAKVLR SIYEF LSAEKREFRFQLRGVA FVMVEDGWKLLKPEE  
VVINLEYEADFKPYLYKLPLELGT FHQLFKH LGTEDIISTKQYVEVLSRIFKSSEGKQ  
LDPNEMRTVKRVVSGLFKSLQND SVKVRSDLENARDL ALYLP S QDGKLVKSS ILV FDD  
APHYKSRIQGNIGVQMLVDLSQC YLGKD HGFHTKLIMLFPQKLRP RLLSSILEEQLDE  
ETPKVCQFGALCSLQGRLQ LLLSSEQFITGLIRIMKHENDNAFLANEEKAIRLCKALR  
EGLKVSCFEKLQTTLRVKGFNP IPHSRSETFAFLKRGNAVILLYIQHSDSKDINFL  
ALA MTLKSATDNLIS DTSYLIAMI LGCNDIYRISEKLDLSLGVKYDSSEPSKLELPMPGT  
PIPAEIHYTLLMDPMNVFYPGEYVG YLVDAEGGDIYGSYQPTTYAIIVQEVEREDAD  
NTSFLGKIYQIDIGYSEYKIVSSL DLYKFSRPDESSQRDSAPTTPTSPTEFLTPGLR  
SIPPLFSGKESHKSPSTKHHSPRKLKVNALPEILKEVTSVVEQAWKLPE SERKKIIRR  
LYLKWHPDKNPENHDIANEVFKHLQNEINRLEKQAF LDQNADRASRRTFSTSASRFQS  
DKYSFQRFYTSWNQEATSHKSERQQSKEKC PPSAGQTY SQRFFVPPTFKSVG NPVEA  
RRWLRQARANFS AARNDLHK NANEWVCFKCYLSTKLALIAADYAVRGKSDKDV KPTAL  
AQKIEEYSQQLEG LTNDVHTLEAYGVDSLKTRYPDLLPFPQI PNDRFTSEVAMRVMEC  
TACIIIKLENFIQQKV"

Figure 8C

BASE COUNT 3599 a 2281 c 2387 g 3226 t

ORIGIN

1 atgaatacat tctggcctgg tcgagagttg gtggttcagt ggtatccatt tagtgaagac  
61 aaacgtcacc catcccttc atggcttaag atggtttggaa agaatctcta tatacatttc  
121 tcggaagatt tgactttatt tgatgagatg ccacttatcc ctagaactct actgaatgag  
181 gaccagacgt gtgtgaaact catcagáctc aggatccat cagtagtcat ttttagatgat  
241 gaaactgaag ctcagcttcc agaattctta gcagatattg tacaaaaact tggagggatt  
301 gtcctgaaaaa gactagatac ctctattcag catccacttg ttaaaaaaaata cattcattcc  
361 ccactcccga gtgctatttt gcagataatg gagaagatac ctctacagaa gttgttaat  
421 caaatagcat cattacttcc aacccacaaa gatgctctaa ggaagttttt ggccagctt  
481 actgatacca gtgaaaaaaga gaaaagaata attcaagaat tgacaatatt caaaagaatt  
541 aatcactcat cagatcaagg gatttcctct tacacaaaat taaaaggatg taaagtttg  
601 gatcataccg ccaagcttcc aacagatcta cggctatcag tttcagtaat agatagtatg  
661 gatgaagcca ccattcggtt ggcaaacatg ttgaaaattt gaaaattgaa gactacaacg  
721 tggtaaagt ttgtttaaa agatattgga aatgcatttt atacacagga agaggtaaaca  
781 caacttatgc ttggatcct tgagaatcta tcctctctt aaaaatgagaa ttcaaatgtg  
841 cttgatttgtt taatgccact aaaattcatt catatgtccc agggacatgt ggttagcagct  
901 ggtgatctct ttgatcctga tatagaagta ctaaggatc tctttataa tgaagaagaa  
961 gcttgttcc cacctacaat ttacactca ccagatatcc ttcaactctt gagacagatt  
1021 ggctaaaaaa atgaatccag tctaaaaagaa aaagatgtt tacaagtggc aaaaaaaatt  
1081 gaagcttac aggtcagttc ctgtcagaat caggatgttc tcatgaagaa agccaaaaca  
1141 ctcttactgg tcttgaataa aaaccagaca ctctgcagt cttctgaagg gaagatggca  
1201 ttgaagaaaaa tcaaataatgggt tccagcctgc aaggaaagac ctccaaatta tccccgttcc  
1261 ttagtcttgg aaggggatct ctgtatctt tgcacccctc cagatatgtg tgatgcggca  
1321 catgcagttc tagtaggctc ctcacttcctt cttgttggaaa gtgtccatgtt gaacctggag  
1381 caggcgctca gcatcttcac aaagccact atcaatgctg tcttaaaaca ctttaaaact  
1441 gttgttactt ggtataacttc aaaaacctt agtgcataatgattactatca gttccaaacat  
1501 attttgccttgg aaatttatgg gttcatgcat gatcatctga gtgaaggaa ggattctttt  
1561 aaagccttga agttccatg gttttggact ggcaaaaact tttgtcctct tgcccgaggct  
1621 gtgataaaagc caacccatga tctggatctt cagccttatt tatataatgt gcctaaaacc  
1681 atggcaaaat tccaccagct gttcaaggct tggctcaaa tagaagatgtt gacatcagat  
1741 catatttcca tggtcattca gaaagttt ctcaaaatg accaggagtt gagtgaagaa  
1801 gaaagtaaac aaaatcttca tctcatgtt aatattatga gatggctcta tagcaatcag  
1861 attccagcaa gccctaatac accagttctt atttacaca gcagaaatcc ttccaaactt  
1921 gtcatgaagc caattcatga atgctttat tgcacatca aagttgatga cctcaatgac  
1981 ttgcttggaa attcagttgg accaatttac ttggatcatg aagatatacc catggaaact  
2041 gcagaatggc taaaagttcc gtgccttagt acaagactga tcaatcctga aaacatgggg  
2101 ttggagcactt cagggcaaaag agacccctt actgttggaa ttaaaaat tttggaaagaa  
2161 tacccttccg tgcacatattttttttt ctacttcaaa atgctgtatg tgcaaatgcc  
2221 acagaatgca gttcatgtat tgatatgaga aggaatatgg acatacgga aaatctccctg  
2281 gaccagggaa tggcagcttgc tcatggaccc gctctgtggt cattcaacaa ttctgaattc  
2341 tcagattcag atttctttaaa cataacgagg ttaggagatg ctttaaaatggggagaatgtt  
2401 gacaaggatgg gggaaatttgg tcttggttt aattctgtgtt accacatcac tgacattccc  
2461 atcattatga gcagagaatt tatgataatg tttgatccaa acataaatca tatcagcaaa  
2521 cacattaaag atagatcgaa tcctggaaatc aaaattaatt ggatgttggca gcagaaaaga  
2581 cttaggaatgt tccccaacca gttcaacca ttatagatg tattttggctg tcagttaccc  
2641 ttggcttgg aagcttccat cagctacaat ggaactctt tccgactgtc cttagaaaca  
2701 cagcagggaa caaaaatgttgg tgaagtttgc agtacttgc acaatactgc ggatattttac  
2761 tcccttagtgg atgaatttttgc tctttgtggg cacagacttca tcatttttac tcagatgtt  
2821 aactcgatgtt atttggaaatc ttttttttttggaaatggatgttggaaatgttggatgtt  
2881 acaatcataa ttaaaaaaaa agtttggccccc tccaaagcat tgaatgcacc agtttaatgt  
2941 gttttaaaatggaaatgttggccccc tccaaagcat tgaatgcacc agtttaatgttggatgtt  
3001 acggatgtgc caaagtcatc ttgcatttcc cagatcacatg tggatgttggatgttggatgtt  
3061 tttaggatgttggccccc tccaaagcat tgaatgcacc agtttaatgttggatgttggatgtt  
3121 actctcttttgc aatggctaa atctggccaa tcaaaaaaagc catcagatgttggatgttggatgtt  
3181 aagacagtagt attgttaccat atggcttata tgcacatgttggatgttggatgttggatgtt  
3241 aagttttccat tgaatgttggccccc tccaaagcat tgaatgcacc agtttaatgttggatgtt

Figure 8D

3301 gttctcttgc atgaaaccca ggaacagaag tggaccgtga aaccacacat aggagaagt  
3361 ttttgctatt tacctctacg aatcaaaaaca gggttgc当地 ttcacatcaa tgggtgc当地  
3421 gctgttactt caaataggaa agaaatctgg aagacagata caaaaggctcg atgaaatacc  
3481 acattcatga ggcattgtcat tgtgaaagct tacttacaag ccctcagtgt cttacgggac  
3541 ctagccattt gttgtgagct gactgattat acttactatg cagtgtggcc tgatcctgat  
3601 ctagttcatg atgacttctc tgtgatctgt aaaggattt atgaaagacat tgctcatggg  
3661 aaggggaaagg agttgaccag agtcttctc gatgggtcta tgtgggttcc catgaagaat  
3721 gtgagggttcc ttgatgactc tatacttcaa aggaaagatg ttggttc当地 agccttcaag  
3781 atatttctga agtacctcaa gaaaacagga tccaaaaacc tctgtgctgt tgagcttc当地  
3841 tcttcagtaa aagcaggatt tgaagaggct ggctgttaaagc agatactgct gggaaaataca  
3901 ttttcagaga aacagtttctc ttcaagatc ttcttc当地 atatccagga aattgaagca  
3961 gaacttagag atcctctgtat gaattttgtc ctaaatgaaa aacttgc当地 gttctc当地  
4021 atttctc当地 ttaccccttgc ttgttgc当地 tccttgagg gccatc当地 ggtttgc当地  
4081 tcaagattga tccatc当地 gaaggacgatg gcaaaagttt ttgatactaa agatggaaagg  
4141 ttcccttatg gttccacaca ggattacctc aatccttata tcttgattaa gctcgttc当地  
4201 ttaggc当地 ggaaaagatga tattttgtgg gatgacatgc tagagcgtc agagtctgta  
4261 gctgagatta ataaaaagtga ccatgctgt gcctgcttaa ggagtagtat tctgctaaagc  
4321 cttattgtatg agaagctaaa aataaaggat cctagagcaa aggatttgc tgcaaaaatata  
4381 caaacaattt cttccctccc atttctaaca aagccagcag gttttctt当地 agaatggaaa  
4441 gggaaacagct ttaagcctga aaccatgttt gcagcaactg acatttacac agctgaaatat  
4501 caagatatacg tctgtctttt gcaaccaattt ctaatgaaa attccattt ctttagaggc  
4561 tgggtttagt gtttttggc tggtaaggag tttttgggt tactaaagaa gccaacagtt  
4621 gatctggtaa taaaccaggat gaagcaaggat gcaaaatcag ttgatgatgg cattacattt  
4681 taccaggaaa atatcacca cgttgc当地 aaataccttgc atgaaagc当地 attgc当地  
4741 gaaatggcca aggcaacaat tattgagaag ctaaaggccat tttttttcat tctagttgag  
4801 aatgtatatg ttgagtc当地 aaaggtttct tttcaacttgc actttgaagc agcaccatc  
4861 ctttatc当地 tacctaacaat gtataaaaat aatttccgtg agcttttgc aagtgtgggt  
4921 gtgc当地 agacttgc当地 tgaagactttt gccc当地 tagtctat tgatcaagag  
4981 agaggaaaaaa aacaataaac agaagagaat tttc当地 gccc当地 gacgaaatc当地  
5041 ggc当地 ctggtaa gtc当地 catttgc当地 agaaaagaga caagaatttt gtagggaaa ttatggcaaa  
5101 atattactgc cagacactaa cctgctgtc当地 ctccctgctt当地 agtc当地 ctaatgac  
5161 tggccctgtaa taaaagtaaa ggactccact gtc当地 agtattt gccatgccc当地 cataccccgg  
5221 gaagtagctg taaaacttgg tgc当地 atacca aagagacata aagc当地 ttagatgca  
5281 tccaaatctt gttcacagc tcttagtaca gaattttggc agaaaagaaa actgaccagc  
5341 agaattaaga gcatcttgc当地 tgc当地 ttttgc当地 tc当地 agatgtgaa agagcttctt  
5401 caaaatgctg atgatgcaaa ggccacacagat atctgctttt tggttgc当地 tagacagcat  
5461 cctgttgc当地 gaatatttgc当地 tgataatgtgg gccc当地 actgtgtgat  
5521 tacaacaacc agccatttac agaagatgtat gtttagaggaa ttc当地 agatct tggaaaggc  
5581 accaaagaag ggaatccctt当地 caaaaacagga cattatggaa tc当地 gatttcaat tccgtt当地  
5641 catattacag actgccc当地 ttttatttgc当地 ggcaatgaca tc当地 tgggtt当地 ttttgc当地  
5701 catgccagat atgc当地 caccagg agccacatca gtttagccctg gacgcatgtt tagagatttgc  
5761 gatgc当地 agacttgc当地 tt当地 agacat gttctc当地 gttctagatc tttacttggg aaaccactt  
5821 aaactggaca attgtacaat gtttagattt cctctgc当地 atgc当地 agat ggc当地 cacaagtt  
5881 tc当地 agaaatctt当地 ct当地 cc当地 gtttgc当地 atcatc当地 gagat ggc当地 cacaagtt  
5941 cggctgtatg gggc当地 gagact tcttaatgtttt ct当地 caaccaca tggagaaaat atcttatttgc  
6001 gaaatagata aggccacagg aggtctgaat gtc当地 ttttgc当地 ctttgc当地 ctttgc当地  
6061 gatggagacc gattgaaaag gaagcaattt caccgc当地 ct当地 taatttgc当地 ttttactt  
6121 aagagacagc tcaaggacat accaggatca caaaatataatcctt acactatggc tactgaggat  
6181 tctgaaggaa atctgaccac atggctc当地 tgc当地 atatgat caggattttc aagtatggaa  
6241 aaagtatcca agagtgtat atc当地 agctc当地 aagaaccaag atatcaccctt tttccacgt  
6301 ggtggagtag cagcctgc当地 tactcacaat tataaaaagc cccacacagc cttctgctt  
6361 ctgc当地 ct当地 tt当地 ggagac agggctgc当地 tttcatgtgc当地 atggccactt tgctctagat  
6421 tc当地 agccagaa gaaacttgc当地 gctgtatgat aatgggggttgc当地 tggttgc当地 tgactggaaat  
6481 aatagttaa tgacagcattt aatagcaccctt gcatatgttgc当地 agttactaat ccagttaaa  
6541 aaacggtattt tccctgggtt当地 tgacccaaaca ttatc当地 agtacacac acccattcat  
6601 gtc当地 taaagg acacattaaa gaagtttgc当地 tc当地 tt当地 ctttgc当地 ctttgc当地  
6661 cagccggact tatattgc当地 agtaaaagca ctttgc当地 agtacatgc当地 agacatgaag

Figure 8E

Figure 8F

10141 aggatcagtg agaagcttga cagtttaggg gtgaaatacg actcctctga gccatcaaaa  
10201 ctggactcc ccatgcctgg cacaccaata cccgctgaga tccattacac actacttatg  
10261 gatccaatga atgttttta tcctgggaa tatgttgggtt accttgtgga tgctgaaggt  
10321 ggtgatatct atgggtcata ccagccaaca tacacatacg caattattgt gcaagaagtt  
10381 gaaagagaag atgctgacaa tactagttc ttaggaaaga tctatcagat cgatattggc  
10441 tacagtgaat ataagatagt cagctctt gatctgtaca agtttcag gcctgatgaa  
10501 agctcccaa acagagacag tgctcccacc acaccaacaa gccccaccga attcctgact  
10561 cctggctctga gaagcatccc tcctctttc tctggcaagg agagccacaa gtctccctcc  
10621 accaaacacc attcccccaag aaagctcaag gtgaatgctt taccagaaat cttaaaagaa  
10681 gtgacatcag tggtgagca agcttgaag cttccagaat cagagcgaa aaagatcatt  
10741 agacgcttgt atttgaagtg gcaccctgac aaaaatccag aaaatcatga tattgctaatt  
10801 gaagtgttca agacacctgca gaatgaaatc aacagattag aaaaacaggc ttttctggat  
10861 caaaatgcag acagagcttc aagaagaaca ttttcaacct ctgcacatctcg atttcagtc  
10921 gacaagatact catttcaaag attttacact tcgttggaaatc aagaagccac aagtctaaaa  
10981 tctgaaaggc aacagcaaag caaagagaaa tgccctcctt ctgctggaca gacataactct  
11041 caaaggttct ttgttccctcc cacctcaag tcagtggca atccagtggaa agcccgaggaa  
11101 tggtaagac aagccagagc aaacttctca gctgccagga atgaccttca caaaaatgcc  
11161 aatgaatggg tggcttcaa gtgttacatt tccaccaagc tggctttgat tgcagccgac  
11221 tatgctgtca gggggaaatc tgataaaagat gtaaagccaa ctgcacattgc aaaaaagata  
11281 gaggagatca gtcagcagct ggaaggactg acaaacgatg tgcacacatt ggaagctt  
11341 ggtgttagaca gcttggaaac aaggtaccct gatttgcttc cttttccgca gattcccaat  
11401 gacaggttca catctgaggt tgccatgagg gtgatggaaat gcactgcctg tatcatcata  
11461 aaacttgaaa attttataca acagaaggtg tga

//

Figure 8G

LOCUS AF193556 12793 bp DNA PRI

DEFINITION Homo sapiens sacsin (SACS) gene, complete cds.

ACCESSION AF193556 VERSION AF193556.1 GI:6907041 KEYWORDS .

SOURCE human.

ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 12793) AUTHORS Engert,J.C., Berube,P., Mercier,J., Dore,C., Lepage,P., Ge,B., Bouchard,J.P., Mathieu,J., Melancon,S.B., Schalling,M., Lander,E.S., Morgan,K., Hudson,T.J. and Richter,A. TITLE ARSACS, a spastic ataxia common in northeastern Quebec, is caused by mutations in a new gene encoding an 11.5-kb ORF JOURNAL Nat. Genet. 24 (2), 120-125 (2000)  
MEDLINE 20120709

REFERENCE 2 (bases 1 to 12793) AUTHORS Engert,J.C., Berube,P., Dore,C., Lepage,P., Ge,B., Hudson,T.J. and Richter,A. TITLE Direct Submission JOURNAL Submitted (08-OCT-1999) Genome Centre, Montreal General Hospital, 1650 Cedar Ave., Montreal, QC H3G 1A4, Canada FEATURES Location/Qualifiers source 1..12793 /organism="Homo sapiens" /db\_xref="taxon:9606" /chromosome="13" /map="between D13S232 and D13S292" mRNA 1..12793 /gene="SACS" /product="sacsin" gene 1..12793 /gene="SACS" CDS 77..11566 /gene="SACS" /note="molecular chaperone" /codon\_start=1 /product="sacsin" /protein\_id="AAF31262.1" /db\_xref="GI:6907042" /translation="

MNTFWPGRELIVQWYPFDENRNHPSVSLKMWKNLYIHFSEDL  
TLFDEMPLIPRTILEEGQTCVELIRLRIPSLVILDDESEAQLPEFLADIVQKLGGFVL  
KKLDASIQHPLIKKYIHSPPLPSAVLQIMEKMPHQKLCNQITSLLPTHKDALRKFLASL  
TDSSEKEKRIIQELAIFKRINHSSDQGISSYTKLGCKVLHHTAKLPADLRLSISVID  
SSDEATIRLANMLKIEQLKTTSCLKVLKDIEENAFYSHEEVTLQMLWVLENLSSLKNE  
NPNVLEWLPLKFIQISQEQMVSAGELFDPDIEVLKDFCNEEGTYFPPSVFTSPDIL  
HSLRQIGLKNEASLKEKDVVQVAKKIEALQVGACPDQDVLLKKAKTLLLVLNKNHTLL  
QSSEGKMTLKKIKWVPACKERPPNYPGSLVWKGDLCNLCAAPPDMCDVGHAILIGSSLP  
LVESIHVNLEKALGIFTKPSLSAVLKHKFIVVDWYSSKTFSDEDYYQFQHILLEIYGF  
MHDHLNEGKDSFRALKFPWTGKFCPLAQAVIKPIHDDLQPYLHNVPKTMKFHQ  
LFKVCGSIEELTSDHISMVIQKIYLKSDQDLSEQESKQNLHMLNIIRWLYSNQIPAS  
PNTPVPIHHSKNPSKLIKMKPIHECCYCDIKVDDLNDLLEDSEVEPIILVHEDIPMKTAE  
WLKVPCLOUDLINPENMGFEQSGQREPLTVRIKNILEYPSVSDIFKELLQNADDANA  
TECSFLIDMRNMDIRENLLDPGMAACHGPALWSFNNSQFSDFVNITRLGESLKRG  
EVDKVGKFGLGFNSVYHITDIPIIMSREFMIMFDPNINHISKHIDKSNPQIKINWSK  
QQKRLRKFPNQFKPFIDVFGCQLPLTVEAPSYNGTLFRLSFRQQEAKVSEVSSTCY  
NTADIYSLVDEFSLCGHRLIIFTQSVKSMYLKYLKIEETNPSLAQDTVIICKKSCSSK  
ALNTPVLSVLKEAAKLMKTCSSSNKLPSPDEPKSSCILQITVEEFHHVFRRIAQLQSP  
LFRGPDDDPAAFLFEMAKSGQSKKPSDELSQKTVECTTWLLCTCMDTGEALKFSLSESG  
RRGLVPGAVGVQLSEIQDQKWTVKPHIGEVFCYLPLRIKTRGLPVHINGCFAVTSNR  
KEIWKTDTKGWRNTTFMRHVIVKAYLQVLSVLRLATSGELMDTYYAVWPDPDLVHD  
DFSVICQGFYEDIAHGKGKELTKVFDGSTWVSMKVNVRFLDDSIKRRDVGSAAFKIF  
LKYLKKTGSKNLCAVELPSSVKGFFEEAGCKQILLENTFSEKQFFSEVFFPNIQEIEA  
ELRDPLMIFVNLNEKVDESGVLRVTPCIPCSLEGHPLVLPSPRLIHPEGRAVAKLFDIKD  
GRFPYGSTQDYLNPIIILKLVQLGMAKDDILWDDMLERAWSVAEINKSDHVAACLRSS  
ILLSLIDEKLKIRDPRAKDFAAKYQTIRFLPFLTKPAGFSLDWKGNFSFKPETMFAATD  
LYTAEHQDIVCLLQPILNENSHSFRGCGSVSLAVKEFLGLKKPTVVLVINQLKEVAK  
SVDDGITLYQENITNACYKYLHEALMQNEITKMSIIDKLKPFSFILVENAYVDSEKVS  
FHLNFEAAPLYQLPNKYKNNFRELFETVGVRQSCTVEDFALVLESIDQERGTQITE  
ENFQLCRRIIISEGIWSLIREKKQEFCEKNYGKILLPDTNLMLLPAKSLCYNDCPWIKV

Figure 9A

KDTTVKYCHADI PREVAVKLGAVPKRHAKALERYASNCFTTLGTEFGQKEKLTSRIKS  
ILNAYPSEKEMLKELLQNADDAKATEICFVDPQRHPVDRIFDDKWAPLQGPALCVYN  
NQPFTEDDVVRGIQNLGKGTKEGNP YKTGQYIGFNSVYHITDCPSFISGNDILCIFDP  
HARYAPGATSISPGRMFRLDADFR TQFSVLDLYLGTHFKLDNCTMFRFPLRNAEMA  
KVSEISSVPASDRMVQNLLDKLRSRGAE MLMFLNHMEKISICEIDKSTGALNVLYSVK  
GKITDGDR LKR KQFH ASVIDS VT KKRQLKDIPVQQITYTMDTEDSEGNLTTWLICNRS  
GFSSMEKVSKSVISAHKNQDITLFPRGGVAACITHNYKKPHRAFCFLPLSLETGLPFH  
VN GHFALDSARRNLWRDDNGVGVRSDWNNSLMTALIAPAYVELLIQLKKRYFP GSDPT  
LSVLQNTPIHVVKDTLKKFLSFFFVNRLDLQPDLYCLVKALYNCIHEDMKRLLPVRA  
PNIDGSDLHS AVI ITWINMSTS NKR PFFDNLLQDELOHLKNADYNITTRKTV AENVY  
RLKHLL EIGFNLVY NCDETANLYHCLIDADIPVSYVTPADIRSF LMTFSSPDTNCHI  
GKLP CRLQQTNLKL FHSLKLLVDYCFKDAEENEIEVEGLPLLITLDSVLQTFDAKRPK  
FLTTYHELI PSRKDLFMNTLYLKYSNILLNCKVAKVFDI SSFA D LSSVLPREYKTKS  
CTKWDNFAS ESWLKNAWHFISESVVKEDQEETKPTFDIVVDTLKD WALLPGTKFTV  
SANQLVVPEGDVLLPLSLMHI A VFPNAQSDKVFHALMKAGCIQLALNKICSKDSA FVP  
LLSCHTANIESPTSILK ALHYMVQTSTFRAEKLVENDFEALLMYFNCNLNHLMSQDDI  
KILKSLPCYKSISGRYVSIGKFGTCYVLTKSIPS AVEKWTQSSSAFLEEKIHLKEL  
YEVIGCVPVDDLEVYLKHL PKIENL S YDAKLEH LIY LKNRLSSA EELSEIKEQLFEK  
LESLLIIHDANSRLKQAHFYDRTVRVFEVMLPEKLFIPNDFKKLEQLIKPKNHVT  
MTSWVEFLRNIGLKYI LSQQQLLQFAKEISVRANTENWSKETLQNTVDILLHHIFQER  
MDLLSGNFLKELSLIPFLCPERA PAEFIRFHPQYQEVNGTLPLIKFNGAQVNPKFQC  
DVLQLLWTSCPILPEKATPLS I KEQEGSDLGPQE QLEQV L MNVNLDPP LDKV INNC  
RNICNITTLD EEMVKTRAKVLR SIYEFLSAEKR EFRFQLRGVAFVMVEDGWKLLKPEE  
V V INLEYESDFKP YLYKL PLEL GTFHQLFKH LGT EDI I STKQYVEVLSRIFKNSEGKQ  
LDPNEMRTVKRVVSGLFRSLQND SVKVRSDLEN VRDLALYLP S QDGR LVKSSIL VFDD  
A PHYKSRIQGNIGVQMLV DLSQCYLGKD HGFHTK LIMLFPQKLRP RLLSSILEEQLDE  
ETPKVCQFGALCSLQGRLQ LSSSEQFITGLIRIMKHENDNAFLANEKAIRLCKALR  
EGLKVSCFEKLQTTLRVKG FNPIPHSRSETFAFLKRG FNAVILLYIQHSDSKD INFLL  
ALAMTLKSATDNLIS DTSYI AMLG CNDIYRIGEKLDSLGV KYDSSEPSKLELPMPGT  
PIP AEI HYTLLMDPMNVFPGEYVG YLVDAEGGDIYGSYQPTTYAIIVQEVEREDAD  
NSSFLGKIYQIDIGYSEYKIVSSLDLYKFSRPEESSQSRDSAPSTPTSPTEFLTPGLR  
SIPPLFSGRESHKTSSKHQSPKKLKVN SLPEILKEV TS VVEQAWKLPESERKKIIRRL  
YLKWHPDKNPENHDIA NEVFKHLQNEINR LEQAFQFLDQNADRASRRTFSTSASRFQSD  
KYSFQR FYT SWNQEATSHKSERQQQNKEKCP PSAGQTY SQRFFV PPTFKS VGNPVEAR  
RWLRQARANFSAARNDLHKNA NEWVCFKC YLSTKLALIAADYAVRGKSDKDVKPTALA  
QKIEEYSQQLEG LTN DVHTLEAYGVDSLKTRYPDLLPFPQIPNDRFTSEVAMRV MECT  
ACIIKLENFMQQKV"

BASE COUNT 4163 a 2256 c 2487 g 3887 t ORIGIN

atgatttaca ggaagaccat gtactcagct gcagcttcta aatccagaac gatttgac ag  
tcttatcaag gaagtaatga atacattctg gcctggcaga gaattgattt ttcaatggta  
tccattttagt gaaaacagaa atcaccatc tgtttcatgg cttaaatggatgg tttggaaaa  
tctttatata catttttcag aggatgttgc ttattttatggat gagatgccac ttatccccag  
aactatacta gaggaaggc agacatgtgt ggaactcatt agactcaggta ttccatcgat  
agtcatttta gacgatgaat ctgaagcaca gcttccagaa ttttttagcag acattgtaca  
aaaacttggaa gggtttgtcc ttaaaaaattt agatgcattt atacaacatc cgcttattaa  
aaaatatattt cattcaccat taccaagtgc tgtttgcag ataatggaga agatgccatt  
gcagaaaattt tgtaatcaaa taacttcgc acttccaaca cacaaggatg ccctgaggaa  
gttcttgct agtttaaccg atagcagtga gaaagagaaa agaattttc aagaatttggc  
aatattcaag cgcatttaacc attcttctga tcaggaaattt tccttcttata caaaatttggaa  
aggttggaaa gtc ttacacc atactgcca actcccgac gatctgcgac tttcttattt  
agtaatagac agtagtgc atagcactat tcgtctggca aacatgttga aaatagaaca  
gttaaagacc actagcgtct aaaaatggat attggaaaatg cattttattt  
acatgaagag gtaacacacgc ttatgtttagt ggtc ttgag aatctatctt ctcttaaaaa

Figure 9B

tgagaatcca aatgtgcttg agtggtaac accattaaaa ttcatccaga tatcacagga  
acagatggta tcagctggtg aactcttga ccctgatata gaagtactaa aggatcttt  
ttgtaatgaa gaaggaacct atttcccacc ctcagtttt acctcaccag atattctca  
ctcctaaga cagattgggt taaaaaacga agccagtctc aaagaaaagg atgttgca  
agtggcaaaa aaaattgaag ctttacaggt cggtgcttgt cctgatcaag atgttcttct  
gaagaaagcc aaaaccctct tactggttt aaataagaat cacacactgt tgcaatcatc  
tgaaggaaag atgacattga agaaaataaa atgggttcca gcctgcaagg aaaggcctcc  
aaattatcca ggctcttgg tctggaaagg agatctctgt aatctctgtg caccaccaga  
tatgtgtgat gttaggccatg caattctcat tggctctca cttccctctt ttgaaagtat  
ccatgtaaac ctggaaaaag cattagggat cttcacaaaaa cctagcctta gtgctgtctt  
aaaacacttt aaaattgtt ttgattggta ttcttcaaaaa accttttagt atgaagacta  
ctatcaattc cagcatattt tgctttagat ttacgattc atgcatgatc atctaaatga  
aggaaagat tcttttagag ctttcaaattt tccatgggtt tggactggca aaaagtttg  
tccacttgcc caggctgtga ttaaaccat ccatgatctt gaccttcagc cttatttgca  
taatgtacct aaaaccatgg caaaattcca ccaactattt aaggctgtg gttcaataga  
ggagttgaca tcagatcata ttccatggt tattcagaag atatatctca aaagtgacca  
agatctcagt gaacaagaaa gcaaacaaaa tcttcatctt atgttgaata ttatcagatg  
gctgtatagc aatcagattc cagcaagccc caacacacca gttccatac atcatagcaa  
aaatccttctt aaacttatca tgaagccat tcacgaatgc ttttattgtg acatcaaagt  
tgatgacctt aatgacttac ttgaagattc tgtggAACCA atcattttgg tgcatgagga  
cataccccatg aaaactgcag aatggctaaa agttccatgc cttagttacaa gactgataaa  
tcctgaaaaac atgggatttg agcagtcagg acaaagagag ccacttactg taagaattaa  
aaatattctg gaagaatacc ctcagtgtc agatattttt aaagaactac ttcaaaacgc  
tgatgatgca aatgcaacag aatgcagttt cttgattgt atgagaagaa atatggacat  
aagagagaat ctcctagacc cagggatggc agcttgcattt ggacctgc ttgtgtcatt  
caacaattctt caattctcattt ctttgcattt ttgttgcattt gagaatctt  
aaaaagggga gaagttgaca aagttggaaa atttggctt ggatttaattt ctgtgtacca  
tatcactgac attccatca ttatgatcg ggaattcatg ataatgttcg atccaaacat  
aaatcatatc agtaaacaca ttaaagacaa atccaatcctt gggatcaaaa ttaattggag  
taaacaacag aaaagactta gaaaatttcc taatcagtcc aaccattta tagatgtatt  
tggctgtcag ttaccttga ctgttagaagc accttacagc tataatggaa ccctttccg  
actgtccctt agaactcaac agaagcaaa agtggatgaa gtttagtagt cgtgcacaa  
tacagcagat atttattctc ttgtggatgtt atttagtctc ttgtggacaca ggcttatcat  
tttcaactcag agttaaagt caatgtattt gaagtacttg aaaatttggg aaaccaaccc  
cagtttagca caagatacag taataattaa aaaaaatcc tgctttcca aagcattgaa  
cacacctgtc ttaagtgttt taaaagggc tgcttaagctc atgaagactt gcagcagcag  
taataaaaag cttccactgt atgaacccaa gtcatcttc attcttcaga tcacagtgg  
agaatttcac catgtttca gaaggattgc tgatttacag tcgcccattt tttagggtcc  
agatgatgac ccagctgtc tctttgaaat ggcttaatgc gccaatcaa aaaagccatc  
agatgagttt tcacagaaaa cagtagatgtt taccacgtgg cttctgtgtt cttgcatt  
cacaggagag gctctgaatg ttccctgag tgagatggaa agaagacttag gactgggtcc  
atgtggggca gtaggatgtt agctgtcaga aatccaggac cagaagtggaa cagtgaaacc  
acacattggaa gaggtgtttt gctatttacc ttacatcaaaa aaaacaggt tgccagttca  
tatcaatggg tgctttgtt ttacatcaaaa tagggaaatggaaa cagataaaaa  
aggacgatgg aataccacgt tcatgagaca tggttattgtt aagcttact tacaggtact  
gagttgtctt cgggacactgg ccacttagtgg ggagctaatg gattatactt actatgcagt  
atggcccgat cctgattttt ttcattatgtt ttgttgcattt atttggccaaat gattttatgt  
agatatagtt catggaaaag gggaaaactt gaccaaaatgc ttctctgtt gatctactt  
ggtttccatg aagaacgtaa gatttcttgc tgactctata cttaaaagaa gagatgttgg  
ttcagcagcc ttcaagatattttt gtttgcattt cctcaagaag actgggtccca aaaaccttt  
tgctgttgc ttcccttcc cgtttttttt aggatttgc gaaatgggtcc gcaacagat  
actacttgaa aacacattttt cagagaaaca gtttttttgc gaaatgtttt ttccaaatat  
tcaagaaattt gaagcagaac tttagagatcc tttaatgtatc ttgttcttcaat atgaaaaatgt  
tgatgagttc tcgggatgttcc ttctgttac tccatgttattt cttgttcc tggagggggca

Figure 9C

tcccttggtt ttgccccatcaa gattgatcca ccccgaaagga cgagggtgca agttatttga  
tattaaagat ggagattcc ctatgggtc tactcaggt tatctcaatc ctattattt  
gattaaacta gttcagttag gtatggaaa agatgatatt ttatggatg atatgctaga  
acgtgcagtg tcagtagctg aaattaataa aagtgatcat gttgtgcat gcctaagaag  
tagtatctt ttgagtcctt tcgatgagaa actaaaaata aggatccctt gagcaaagga  
ttttgctgca aaatatcaaa caatccgctt cttccattt ctgacaacac cagcaggtt  
ttcttggac tgaaaaggca acagtttaa gcctgaaacc atggttgcag caactgacct  
ttatacagct gaacatcaag atatagttt tctttgca ccaattctaa atgaaaattt  
ccattcttt agaggttgt gttcagtgatc attggctgtt aaagagttt tgggattact  
caagaagcca acagttgatc tggttataaa ccaattgaaa gaagtagcaa aatcagttga  
tgcgttggatt acactgtacc aggagaatatt cacaatgtc tgctacaat accttcatga  
agccttgatc caaaatgaaa tcactaagat gtcaattt gataagttt aaccctttag  
cttcattctt gttgagaatg catatgttgc ctcagaaaag gtttctttc attttaattt  
tgaggcggca ccataccctt atcagttgcc taataagtat aaaaataatt tccgcgaact  
ttttgaaaacc gtgggtgtg ggcagtcatg cactgttga gatttgctc ttgttttgg  
atctattgtat caagaaagag gaacaaagca aataacagaa gagaattttc agcttggcc  
acgataatc agtgaaggaa tatggagtct cattagagaa aagaaacaag aattttgtga  
aaaaattt ggcagatatt tattggcaga tactaatctt atgccttcctc ctgctaaatc  
gttatgtac aatgattgccc ttggataaa agtaaaggat accactgtaa aatattgtca  
tgctgacata cccagggaaag tagcgtttaaa actaggagca gtcggaaagc gacacaaagc  
cttagaaaga tatgcatcca atgtctgtt tacaacactt ggcacagaat ttggcagaa  
agaaaaattt accagcagaa ttaagagcat ccttaatgca tattttctt ggggggg  
gttggaaagag ctcttcaaa atgctgtatc tgcaaaaggcg acagaaatct  
tgatcctaga cagcatccag ttgatagaat atttgcattt aagtgggccc  
gccagcactt tgggtgtaca acaaccagcc atttacagaa gatgtatc  
aatcttggaaagag aaaggcacga aagaggaaa tccttataaa actggacagt  
attcaattct gtgtatcata tcacagactg cccatcttt atttctggca  
gtgtatctt gatcctcatg ccagatatgc accagggggc acatccatta  
catgtttaga gattttggatc cagattttag gacacagttc tcagatgtt  
tctggaaacc cattttaaac tggataattt cacaatgttca agatttctt  
agaaatggca aaagtttcgg aaatttcgtc tggccagca tcagacagaa  
tctttggac aaactgcgtc cagatgggc agaacttcta atgtttctt  
aaaaatttctt atttgcattt tagataagag tactggagct ctaaatgtc  
aaagggcaaa atcacagatc gagacagatt gaaaaggaaa caatttcatg  
tgatagttt actaaaaaga ggcagctcaa agacatacca gttcaacaaa  
tatggataact gaggactctg aaggaaatct tactacgtt ctaatttgc  
ctttcaagt atggagaaag tatctaaaag tgcataatc gtcacaaga  
tactctttc ccacgtggc gagtagctgc ctgcattact cacaactata  
tagggccccc tggggggc ctctttctt ggagactggg ctgcatttc  
ccactttgca ctggatttcag ccagaaggaa cctgtggcgt gatgataatc  
tcgaagtgc tggataaca gtttaatgc agcattaata gtcctgc  
gctaatacag taaaaaaaac ggtattttcc tgggttctgtt ccaacattat  
gaacacccctt attcatgtt taaaggacac tttaaagaag tttttatcgt  
taaccgtctt gatctacagc cagattata ttgtcttagt aaagacttt  
tcacgaagac atgaaacgtc tttacctgt tggccggc ccaaatatt  
cttgcactct gcaagttataa ttacttggat caatatgtt acttcttata  
attttttgac aatttactac aggttgcattt acaacaccc  
caccacacgc aaaacagtag cagagaatgt ctataggctg  
tggttcaac ttgggttata actgtatc aactgcttact  
tgcagatatt cctgttagt atgtgacccc tgctgatatc  
ttcctctcctt gacactaattt gccatattgg gaagctgc  
tctaaaactt tttcatagtt taaaactttt agttgattt  
aaatgagattt gaaatggagg gattggccct tctcatcaca  
ttttgatgca aaacgaccca agtttcttata aacatatcat  
gaatttgcattt catccccca

Figure 9D

agacttgttt atgaatacat tatatttcaa atatagtaat attttatttg actgtaaagt  
tgcaaaaatgt tttgacattt ccagcttgc tgatttggta tcctctgtgt tgccctcgaga  
atataagacc aaaagttca caaatggaa agacaattt gcaagtggat cttggcttaa  
gaatgcattt cattttatta gtgaatctgt aagtgtgaaa gaagatcagg aagaaacaaa  
accaacattt gacattgtt ttgatactct aaaagactgg gcattgcttc caggaacaaa  
gtttactgtt tcagccaacc agcttgggt tcctgaagga gatgttctgc ttccctctc  
ccttatgcac attgcagtt ttccaaatgc ccagagtat aaagttttc atgctcta  
gaaagccggc tgtattcagc ttgcttggaa caaatctgt tccaaagaca gtcatttgt  
tccttggc tcatgtcaca cagcaaatat agagagcccc acaagcatct tgaaggctct  
acattatatg gtccaaactt caacatttag agcagaaaaa tttagagaaa atgatttga  
ggcactttt atgtatttca actgcaattt gaatcattt atgtcccag atgatataaa  
aattctaaag tcacttccgt gctataaattt catcagtggc cgctatgtaa gcattggaaa  
atttggacaa tgctacgtac ttacaaaaag tatccctca gctgaagtgg agaaatggac  
acaatcatca tcatctgcatt ttcttgaaga aaaaatacac ttaaaaagaac tatacgagg  
gattgggtgt gtacctgtat atgatcttga ggtatattt aaacacctct taccaaaaat  
tgaaaatctc tcttatgtat caaatattttt gcaacttgc tacctttaaga atagattatc  
aagtgttagt gaattatcag agattaagga acaactttt gaaaaactgg aaagtttatt  
gataatccat gatgctaaca gtagactaaa gcaagcaag catttctatg atagaactgt  
gagagttttt gaagttatgc ttccttggaa attgtttattt cctaatttgc tcttttaagaa  
atttggacaa ctataaaaac ccaaaaatca tgttacattt atgacatctt gggtggatt  
cttaaaaat attggactaa aatacataact ttctcagcag cagttgttac agtttgc  
gaaaatcagt gtgagggcta atacagaaaaa ctggtccaaa gaaacattgc aaaatacagt  
tgatattcctt ctgcatcata tatttcaaga acgaatggat ttgttatctg gaaaatttct  
gaaagaacta tcttttaatac catttctatg ttctgagcgg gccccggg aattcattag  
atttcatcctt caatatcaag aggtaatgg aacacttctt cttataaaatg tcaatggagc  
acaggtaaat cccaaattca agcaatgtga tttactccag ctgttatggg catcctggcc  
tatttccca gagaaagcta cacccttaag cattaaagaa caagaaggta gtgaccttgg  
tccacaagaa cagcttgaac aagttttaaa tatgttttaat gttaacctgg atcctcctt  
tgataaggtt atcaataact gcaagaaacat atgcaacata acgacgttgg atgaagaaaat  
ggtttttactt agagcaaaag tcttaaggag catatatggaa ttccctcagtg cagaaaaaaag  
ggaatttctt tttcagttgc gaggggttgc ttttggatg gttagaagatg gttggaaact  
tctgaagcctt gaggaggttgc tttttttttt aaccttattt gttttttttt aaccttattt  
gtacaagcta ctttttagaac ttggcacattt tcaccaggatg ttcaacact taggtactga  
agatattttt tcaactaaggc aatatgttgc aagtgttgc cgcattttttaa aaaattctga  
ggcataacaa tttagatccta atgaaatgcg tacagttaag agagtagttt ctggtctt  
caggagtcta cagaatgttgc cagttttttttt cttttttttt cttttttttt cttttttttt  
tgcccttac ctcccaagcc aggtttttttt tttttttttt tttttttttt tttttttttt  
cgatgcgcca cattataaaa gttagaatcca ggggaaatattt ggtgtgaaa tttttttttt  
tctcagccag tgctacttag ggaaagacca tggatttccactt actaagttga taatgtctt  
tcctcaaaaaa ctttagaccc tttttttttt tttttttttt tttttttttt tttttttttt  
gactcccaaa gtttgcgtt ttggagcgtt gtgttctt caagaaagat tgcaaggat  
cttgcattttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt  
gaacagttca ttacaggact gattttttttt tttttttttt tttttttttt tttttttttt  
tgctttctt gccaatgttgc aaaaagccat aagactttgc aaagccctaa gagaaggatt  
gaaagtatcc tgctttttttt tttttttttt tttttttttt tttttttttt tttttttttt  
tccccacagc agaagtggaaa cttttttttt tttttttttt tttttttttt tttttttttt  
gctctacattt caacatttcag acagttttttt tttttttttt tttttttttt tttttttttt  
tcttaaatca gcaactgaca atttgcattt tttttttttt tttttttttt tttttttttt  
atgcaatgttgc atttgcattt tttttttttt tttttttttt tttttttttt tttttttttt  
ttcggagccca tccaaactgg aacttccat tttttttttt tttttttttt tttttttttt  
tttacactctg cttatggacc caatgttgc tttttttttt tttttttttt tttttttttt  
tggttgcgtt gaaggtgggtt gatctatgg atcataccag ccaacataca catatgc  
tattgtacaa gaagttggaaa gagaagatgc tgacaattt tttttttttt tttttttttt  
tcagatagat attgggttataa gttagatataa aatagttgc tttttttttt tttttttttt  
ttcaagaccc gaggaaagct cttttttttt tttttttttt tttttttttt tttttttttt

Figure 9E

cactgagttc ctccccctcg gcctgagaag cattcctctt ctttctctg gtagagagag  
ccacaagact tcttccaaac atcagtcccc caaaaagctt aaggtaatt ctttaccaga  
aatcttaaaa gaagtgacat ctgtgtgga gcaagcatgg aagcttccag aatcgaaacg  
aaaaaagatt attaggcggt tgtattgaa atggcatctt gacaaaaatc cagagaacca  
tgacattgcc aatgaagttt ttaaacattt gcagaatgaa atcaacagat tagaaaaaca  
ggctttcta gatcaaaatg cagacaggc ctccagacga acatttcaa cctcagcatc  
ccgatttcag tcagacaaat actcattca gagattctat acttcatgga atcaagaacg  
aacgagccat aaatctgaaa gacagcaaca gaacaaagaa aaatcccc cttcagccgg  
acagacttac tctcaaagg tctttgtcc tcccacttc aagtcgggt gcaatccagt  
ggaagcacgc agatggctaa gacaagccag agcaaacttc tcagctgcca ggaatgacct  
tcataaaaat gccaatgagt gggtgtgctt taaatgttac ctttctacca agttagctt  
gattgcagct gactatgctg tgagggaaa gtctgataaa gatgaaaac caactgcact  
tgctcagaaa atagaggaat atagtcagca acttgaagga ctgacaaatg atgttcacac  
attggaagct tatggtag acagttaaa aacaagatac cctgatttgc ttcccttcc  
tcagatccc aatgacagg tcaactctga gttgctatg agggtgatgg aatgtactgc  
ctgtatcata ataaaaactt gaaaattttt gcaacaaaaa gtgtgaagat atttaacgaa  
aaaaaaggta gatcttgaat gtgtttagc acgaataat tgctgtactt cattaagctt  
cattgccaat tagcttagaa ttgttaagca cattgcagat tggtttggaa gaattctgga  
gttggtaga acatgaatac caacggaaaa ctttaactga atctaaaaga aaactattt  
gaagatggg gtagctgca aaatagctgg atggatttga atgattggg tgatacatca  
ttgaactgca ctttatataa ccaaagctt gcagttgtt agataagagt ctatgtatgt  
ctctggtag gatgaagttt atttatgtt ttaacatgg tattttgaa ggagctaing  
aaacactgga catataattt gtttaacat aaggggaaatt aagtctttgt agtctgtcat  
tttttaagt ggtcctctt ggtcgttta tttctcatc agctggctct gatcatgaat  
ttgtttagt tttatgtt gactcgtgca ttaagaaaat ggttaggtat ttaatctca  
ttacttgact aagagtgtga aggttagtact ttttagagtg cactgagtgc actttacatc  
tttatttaaa ttttttttta acatctttag tttacaggct tcctgttga tgaagatagc  
aacggaaaac tcaaaatggt ggcagttctt attaccagg ttttagtattt tttctggaaa  
ctgcttgc aagacaacatt tattaactgt tagaacactt gctttatgtt tttgtgtaca  
tattttccac aaatgttata atttatata gttgggtt gaa caggatgca tcttttttttgc  
tctaaagggtg ctgcagttaa aaaaaaaaaaca accttttctt tcaatatggc atgttagtgga  
gttttttaa cttaaaaac atcaaaaatt gttaaaatca ttgtttagt ttttttttttgc  
taattatcg ctttatattt cccatgaatg atcagaactg acatttaatt tagtagttt  
tcgccatgtc tctttacttt aacatattt ctttgcagaa tgtaaaaggt aatgataatt  
agtttatata agtgtactgg ctgtaaatga tgctaaat atctttagc attaagggt  
tacagaacat gttgaaactt ttttacttt tattggaaat aaggaatgtt tgcacccctca  
cattttatttgc ctt

Figure 9F